

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/839,185

DATE: 05/18/2001
 TIME: 12:52:22

Input Set : A:\Pto.amc
 Output Set: N:\CRF3\05182001\I839185.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:
 5 (i) APPLICANT: Schmidt, Eduard Daniel Leendert
 6 De Vries, Sape Cornelis
 7 Hecht, Valerie France Gabrielle
 9 (ii) TITLE OF INVENTION: Apomixis Conferred By Expression of SERK
 10 Interacting Proteins
 12 (iii) NUMBER OF SEQUENCES: 18
 14 (iv) CORRESPONDENCE ADDRESS:
 15 (A) ADDRESSEE: Syngenta Patent Dept.
 16 (B) STREET: 3054 Cornwallis Road
 17 (C) CITY: RTP
 18 (D) STATE: NC
 19 (E) COUNTRY: USA
 20 (F) ZIP: 27709
 22 (v) COMPUTER READABLE FORM:
 23 (A) MEDIUM TYPE: Floppy disk
 24 (B) COMPUTER: IBM PC compatible
 25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
 28 (vi) CURRENT APPLICATION DATA:
 C--> 29 (A) APPLICATION NUMBER: US/09/839,185
 C--> 30 (B) FILING DATE: 19-Apr-2001
 31 (C) CLASSIFICATION:
 33 (viii) ATTORNEY/AGENT INFORMATION:
 34 (A) NAME: Meigs, J. Timothy
 35 (B) REGISTRATION NUMBER: 38,241
 36 (C) REFERENCE/DOCKET NUMBER: S-30683A
 38 (ix) TELECOMMUNICATION INFORMATION:
 39 (A) TELEPHONE: 919-541-8587
 42 (2) INFORMATION FOR SEQ ID NO: 1:
 44 (i) SEQUENCE CHARACTERISTICS:
 45 (A) LENGTH: 551 base pairs
 46 (B) TYPE: nucleic acid
 47 (C) STRANDEDNESS: double
 48 (D) TOPOLOGY: linear
 50 (ii) MOLECULE TYPE: cDNA to mRNA
 52 (iii) HYPOTHETICAL: NO
 C--> 54 (iv) ANTI-SENSE: NO
 56 (vi) ORIGINAL SOURCE:
 57 (A) ORGANISM: Arabidopsis thaliana
 59 (vii) IMMEDIATE SOURCE:
 60 (B) CLONE: 3A35
 63 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 65 ACGTGTCCTT GGAGGCGGGT CGGGTCAGTC GGGTCAGATA CCAAGGTGCC AAGTGAAGG 60
 67 TTGTGGGATG GATCTAACCA ATGCAAAAGG TTATTACTCG AGACACCGAG TTTGTGGAGT 120
 69 GCACTCTAAA ACACCTAAAG TCACTGTGGC TGGTATCGAA CAGAGGTTTT GTCAACAGTG 180

RAW SEQUENCE LISTING

DATE: 05/18/2001

PATENT APPLICATION: US/09/839,185

TIME: 12:52:22

Input Set : A:\Pto.amc

Output Set: N:\CRF3\05182001\I839185.raw

```

71 CAGCAGGTTT CATCAGCTTC CGGAATTTGA CCTAGAGAAA AGGAGTTGCC GCAGGAGACT 240
73 CGCTGGTCAT AATGAGCGAC GAAGGAAGCC ACAGCCTGCG TCTCTCTCTG TGTTAGCTTC 300
75 TCGTTACGGG AGGATCGCAC CTTCGCTTTA CGAAAATGGT GATGCTGGAA TGAATGGAAG 360
77 CTTTCTTGGG AACCAAGAGA TAGGATGGCC AAGTTCAAGA ACATTGGATA CAAGAGTGAT 420
79 GAGGCGGCCA GTGTCATCAC CGTCATGGCA GATCAATCCA ATGAATGTAT TTAGTCAAGG 480
81 TTCAGTTGGT GGAGGAAGGA CAAGCTTCTC ATCTCCAGAG ATTATGGACA CTAAACTAGA 540
83 GAGCTACAAG G 551

```

85 (2) INFORMATION FOR SEQ ID NO: 2:

87 (i) SEQUENCE CHARACTERISTICS:

88 (A) LENGTH: 375 amino acids

89 (B) TYPE: amino acid

90 (C) STRANDEDNESS: single

91 (D) TOPOLOGY: linear

93 (ii) MOLECULE TYPE: protein

95 (iii) HYPOTHETICAL: NO

C--> 97 (iv) ANTI-SENSE: NO

99 (vi) ORIGINAL SOURCE:

100 (A) ORGANISM: Arabidopsis thaliana

102 (vii) IMMEDIATE SOURCE:

103 (B) CLONE: 3A35

106 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

108 Met Glu Met Gly Ser Asn Ser Gly Pro Gly His Gly Pro Gly Gln Ala
109 1 5 10 15
111 Glu Ser Gly Gly Ser Ser Thr Glu Ser Ser Ser Phe Ser Gly Gly Leu
112 20 25 30
114 Met Phe Gly Gln Lys Ile Tyr Phe Glu Asp Gly Gly Gly Ser Gly
115 35 40 45
117 Ser Ser Ser Ser Gly Gly Arg Ser Asn Arg Arg Val Arg Gly Gly Gly
118 50 55 60
120 Ser Gly Gln Ser Gly Gln Ile Pro Arg Cys Gln Val Glu Gly Cys Gly
121 65 70 75 80
123 Met Asp Leu Thr Asn Ala Lys Gly Tyr Tyr Ser Arg His Arg Val Cys
124 85 90 95
126 Gly Val His Ser Lys Thr Pro Lys Val Thr Val Ala Gly Ile Glu Gln
127 100 105 110
129 Arg Phe Cys Gln Gln Cys Ser Arg Phe His Gln Leu Pro Glu Phe Asp
130 115 120 125
132 Leu Glu Lys Arg Ser Cys Arg Arg Arg Leu Ala Gly His Asn Glu Arg
133 130 135 140
135 Arg Arg Lys Pro Gln Pro Ala Ser Leu Ser Val Leu Ala Ser Arg Tyr
136 145 150 155 160
138 Gly Arg Ile Ala Pro Ser Leu Tyr Glu Asn Gly Asp Ala Gly Met Asn
139 165 170 175
141 Gly Ser Phe Leu Gly Asn Gln Glu Ile Gly Trp Pro Ser Ser Arg Thr
142 180 185 190
144 Leu Asp Thr Arg Val Met Arg Arg Pro Val Ser Ser Pro Ser Trp Gln
145 195 200 205
147 Ile Asn Pro Met Asn Val Phe Ser Gln Gly Ser Val Gly Gly Gly Arg
148 210 215 220

```

RAW SEQUENCE LISTING

DATE: 05/18/2001

PATENT APPLICATION: US/09/839,185

TIME: 12:52:22

Input Set : A:\Pto.amc

Output Set: N:\CRF3\05182001\I839185.raw

```

150   Thr Ser Phe Ser Ser Pro Glu Ile Met Asp Thr Lys Leu Glu Ser Tyr
151       225                230                235                240
153   Lys Gly Ile Gly Asp Ser Asn Cys Ala Leu Ser Leu Leu Ser Asn Pro
154       245                250                255
156   His Gln Pro His Asp Asn Asn Asn Asn Asn Asn Asn Asn Asn Asn
157       260                265                270
159   Asn Asn Asn Thr Trp Arg Ala Ser Ser Gly Phe Gly Pro Met Thr Val
160       275                280                285
162   Thr Met Ala Gln Pro Pro Pro Ala Pro Ser Gln His Gln Tyr Leu Asn
163       290                295                300
165   Pro Pro Trp Val Phe Lys Asp Asn Asp Asn Asp Met Ser Pro Val Leu
166       305                310                315                320
168   Asn Leu Gly Arg Tyr Thr Glu Pro Asp Asn Cys Gln Ile Ser Ser Gly
169       325                330                335
171   Thr Ala Met Gly Glu Phe Glu Leu Ser Asp His His His Gln Ser Arg
172       340                345                350
174   Arg Gln Tyr Met Glu Asp Glu Asn Thr Arg Ala Tyr Asp Ser Ser Ser
175       355                360                365
177   His His Thr Asn Trp Ser Leu
178       370                375

```

180 (2) INFORMATION FOR SEQ ID NO: 3:

182 (i) SEQUENCE CHARACTERISTICS:

183 (A) LENGTH: 859 base pairs

184 (B) TYPE: nucleic acid

185 (C) STRANDEDNESS: double

186 (D) TOPOLOGY: linear

188 (ii) MOLECULE TYPE: cDNA to mRNA

190 (iii) HYPOTHETICAL: NO

C--> 192 (iv) ANTI-SENSE: NO

194 (vi) ORIGINAL SOURCE:

195 (A) ORGANISM: Arabidopsis thaliana

197 (vii) IMMEDIATE SOURCE:

198 (B) CLONE: 3B39

201 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

203 TCAACATTGC TTCCTAACCA GAAATCCACC ATCATCTTCC CACGAATACA ACTTAAAGCT      60
205 TTACCAGAAA ATGGAGGGTC AGAGAACACA ACGCCGGGGT TACTTGAAAG ACAAGGCTAC      120
207 AGTCTCCAAC CTTGTTGAAG AAGAAATGGA GAATGGCATG GATGGAGAAG AGGAGGATGG      180
209 AGGAGACGAA GACAAAAGGA AGAAGGTGAT GGAAAGAGTT AGAGGTCCTA GCACTGACCG      240
211 TGTTCCATCG CGACTGTGCC AGGTCGATAG GTGCACTGTT AATTTGACTG AGGCCAAGCA      300
213 GTATTACCGC AGACACAGAG TATGTGAAGT ACATGCAAAG GCATCTGCTG CGACTGTTGC      360
215 AGGGGTCAGG CAACGCTTTT GTCAACAATG CAGCAGGTTT CATGAGCTAC CAGAGTTTGA      420
217 TGAAGCTAAA AGAAGCTGCA GGAGGCGCTT AGCTGGACAC AATGAGAGGA GGAGGAAGAT      480
219 CTCTGGTGAC AGTTTTGGAG AAGGGTCAGG CCGGAGAGGG TTTAGCGGTC AACTGATCCA      540
221 GACTCAAGAA AGAAACAGGG TAGACAGGAA ACTTCCTATG ACCAACTCAT CATTTAAGGG      600
223 ACCACAGATC AGATAAACCC TCCCGCTCTC TCTCTTCTGT CATCTACATA TGCTCTATCT      660
225 ACACTCTTAT TAGACAAATA ATGGCATCTA ACAATGTCAA GAAAAGTTGG TCATGGTATT      720
227 AAATCCTAGA GGGAAATATA AGTATAAACC TTTAGTCCCC TTTATGCTGT CCTGTAATGA      780
229 ATATCTATCC GGAAATGTAT TCGCATAGTC TTGCGTCTAA TAATGTTTAT TAAAAAATAA      840
231 AAAAAAATAA AAAAAAATAA

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/839,185

DATE: 05/18/2001
TIME: 12:52:22

Input Set : A:\Pto.amc
Output Set: N:\CRF3\05182001\I839185.raw

233 (2) INFORMATION FOR SEQ ID NO: 4:

235 (i) SEQUENCE CHARACTERISTICS:

236 (A) LENGTH: 181 amino acids

237 (B) TYPE: amino acid

238 (C) STRANDEDNESS: single

239 (D) TOPOLOGY: linear

241 (ii) MOLECULE TYPE: protein

243 (iii) HYPOTHETICAL: NO

C--> 245 (iv) ANTI-SENSE: NO

247 (vi) ORIGINAL SOURCE:

248 (A) ORGANISM: Arabidopsis thaliana

250 (vii) IMMEDIATE SOURCE:

251 (B) CLONE: 3B39

254 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

256	Met	Glu	Gly	Gln	Arg	Thr	Gln	Arg	Arg	Gly	Tyr	Leu	Lys	Asp	Lys	Ala
257	1			5						10					15	
259	Thr	Val	Ser	Asn	Leu	Val	Glu	Glu	Glu	Met	Glu	Asn	Gly	Met	Asp	Gly
260				20					25					30		
262	Glu	Glu	Glu	Asp	Gly	Gly	Asp	Glu	Asp	Lys	Arg	Lys	Lys	Val	Met	Glu
263			35				40					45				
265	Arg	Val	Arg	Gly	Pro	Ser	Thr	Asp	Arg	Val	Pro	Ser	Arg	Leu	Cys	Gln
266		50					55					60				
268	Val	Asp	Arg	Cys	Thr	Val	Asn	Leu	Thr	Glu	Ala	Lys	Gln	Tyr	Tyr	Arg
269		65				70					75				80	
271	Arg	His	Arg	Val	Cys	Glu	Val	His	Ala	Lys	Ala	Ser	Ala	Ala	Thr	Val
272				85						90					95	
274	Ala	Gly	Val	Arg	Gln	Arg	Phe	Cys	Gln	Gln	Cys	Ser	Arg	Phe	His	Glu
275			100						105					110		
277	Leu	Pro	Glu	Phe	Asp	Glu	Ala	Lys	Arg	Ser	Cys	Arg	Arg	Arg	Leu	Ala
278			115					120				125				
280	Gly	His	Asn	Glu	Arg	Arg	Arg	Lys	Ile	Ser	Gly	Asp	Ser	Phe	Gly	Glu
281		130						135				140				
283	Gly	Ser	Gly	Arg	Arg	Gly	Phe	Ser	Gly	Gln	Leu	Ile	Gln	Thr	Gln	Glu
284		145			150					155					160	
286	Arg	Asn	Arg	Val	Asp	Arg	Lys	Leu	Pro	Met	Thr	Asn	Ser	Ser	Phe	Lys
287			165					170							175	
289	Gly	Pro	Gln	Ile	Arg											
290			180													

292 (2) INFORMATION FOR SEQ ID NO: 5:

294 (i) SEQUENCE CHARACTERISTICS:

295 (A) LENGTH: 479 base pairs

296 (B) TYPE: nucleic acid

297 (C) STRANDEDNESS: double

298 (D) TOPOLOGY: linear

300 (ii) MOLECULE TYPE: cDNA to mRNA

302 (iii) HYPOTHETICAL: NO

C--> 304 (iv) ANTI-SENSE: NO

306 (vi) ORIGINAL SOURCE:

307 (A) ORGANISM: Arabidopsis thaliana

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/839,185

DATE: 05/18/2001
TIME: 12:52:22

Input Set : A:\Pto.amc
Output Set: N:\CRF3\05182001\I839185.raw

```

309 (vii) IMMEDIATE SOURCE:
310 (B) CLONE: 4B19
313 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
315 AGAAGCAGAA AGGTAAAGCT ACAAGTAGTA GTGGAGTTTG TCAGGTCGAG AGTTGTACCG 60
317 CGGATATGAG CAAAGCCAAA CAGTACCACA AACGACACAA AGTCTGCCAG TTTCATGCCA 120
319 AAGCTCCTCA TGTTCGGATC TCTGGTCTTC ACCAACGTTT CTGCCAACAA TGCAGCAGGT 180
321 TTCACGCGCT CAGTGAGTTT GATGAAGCCA AGCGGAGTTG CAGGAGACGC TTAGCTGGAC 240
323 ACAACGAGAG AAGGCGGAAA AGCACAACG ACTAAAGACG GTGAAACGTG TGAGATCCCG 300
325 GTTTGAAGGT TAATGAAACA GGCTTTGCTT ACTCTCTTCT GTCAGTCTCT TTTAGCTCCT 360
327 TGTAATCCTC TGTGTCTCTG TCTGTTTCTC CATATTACCT GTAATCAAAG CTATCTGCTA 420
329 AACCTACGAC ATGTTAAAT AAATGCATTG AGACTTAAAA AAAAAAAAAA AAAAAAAAAA 479
331 (2) INFORMATION FOR SEQ ID NO: 6:
333 (i) SEQUENCE CHARACTERISTICS:
334 (A) LENGTH: 131 amino acids
335 (B) TYPE: amino acid
336 (C) STRANDEDNESS: single
337 (D) TOPOLOGY: linear
339 (ii) MOLECULE TYPE: protein
341 (iii) HYPOTHETICAL: NO
C--> 343 (iv) ANTI-SENSE: NO
345 (vi) ORIGINAL SOURCE:
346 (A) ORGANISM: Arabidopsis thaliana
348 (vii) IMMEDIATE SOURCE:
349 (B) CLONE: 4B19
352 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
354 Met Ser Met Arg Arg Ser Lys Ala Glu Gly Lys Arg Ser Leu Arg Glu
355 1 5 10 15
357 Leu Ser Glu Glu Glu Glu Glu Glu Thr Glu Asp Glu Asp Thr
358 20 25 30
360 Phe Glu Glu Glu Glu Ala Leu Glu Lys Lys Gln Lys Gly Lys Ala Thr
361 35 40 45
363 Ser Ser Ser Gly Val Cys Gln Val Glu Ser Cys Thr Ala Asp Met Ser
364 50 55 60
366 Lys Ala Lys Gln Tyr His Lys Arg His Lys Val Cys Gln Phe His Ala
367 65 70 75 80
369 Lys Ala Pro His Val Arg Ile Ser Gly Leu His Gln Arg Phe Cys Gln
370 85 90 95
372 Gln Cys Ser Arg Phe His Ala Leu Ser Glu Phe Asp Glu Ala Lys Arg
373 100 105 110
375 Ser Cys Arg Arg Arg Leu Ala Gly His Asn Glu Arg Arg Arg Lys Ser
376 115 120 125
378 Thr Thr Asp
379 130
381 (2) INFORMATION FOR SEQ ID NO: 7:
383 (i) SEQUENCE CHARACTERISTICS:
384 (A) LENGTH: 2682 base pairs
385 (B) TYPE: nucleic acid
386 (C) STRANDEDNESS: double
387 (D) TOPOLOGY: linear

```

VERIFICATION SUMMARY

DATE: 05/18/2001

PATENT APPLICATION: US/09/839,185

TIME: 12:52:23

Input Set : A:\Pto.amc

Output Set: N:\CRF3\05182001\I839185.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:54 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:97 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:192 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:245 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:304 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:343 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:393 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:506 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:689 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:732 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:803 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:848 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:904 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:947 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1024 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1065 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1112 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1135 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]